確率・統計基礎: 仮説検定

森 立平

仮説検定

答が二値の推定問題を考えたいが事前分布を適切に決める方法がない。

- 開発中の薬に効果があるかないか。
- 患者が病気かどうか。
- サイコロに偏りがあるかどうか。

仮説検定

- 対立仮説
  - 開発中の薬に効果がある。
  - 患者が病気である。
  - サイコロに偏りがある。

- 帰無仮説
  - 開発中の薬に効果がない。
  - 患者が病気でない。
  - サイコロに偏りがない。

帰無仮説を採択 or 棄却する。
対立仮説を採択 or 棄却するのではない。

帰無仮説を棄却 $\Longrightarrow$ 対立仮説が正しいと考える。
帰無仮説を採択 $\Longrightarrow$ 何も言えない。

数学 I の仮説検定

コインの仮説検定

コインを 10 回投げたら 9 回表が出た。このコインは偏りが無いと言えるか？

- 帰無仮説: コインに偏りはない
- 対立仮説: コインに偏りはある

帰無仮説が正しいとしてこのような事象が起きる確率を考える。
表が 0,1,9,10 回ちょうど出る確率は $\frac{22}{1024} \approx 2.15\%$ である。
有意水準を 5% とすると、これは有意水準より小さい確率なので帰無仮説を棄却する。コインは偏っていると考える。

数学 B の仮説検定

コインの仮説検定

コインを 10000 回投げたら 5100 回表が出た。このコインは偏りが無いと言えるか？

- 帰無仮説: コインに偏りはない
- 対立仮説: コインに偏りはある

帰無仮説が正しいとしてこのような事象が起きる確率を考える。表が 4900 回以下もしくは 5100 回以上出る確率を標準正規分布を使って近似する。
表が出る回数を表す確率変数を X とすると、$\frac{X-5000}{\sqrt{10000\cdot\frac{1}{4}}}$ は近似的に標準正規分布 $N(0,1)$ に従う (中心極限定理)。
また、$\int_{-1.96}^{1.96}\frac{1}{\sqrt{2\pi}}\mathrm{e}^{-\frac{x^2}{2}}\,\mathrm{d}x \approx 0.95$ である。一方で $\frac{5100-5000}{50} = 2 > 1.96$ である。
よって有意水準を 5% とすると、ここまで偏る事象はこれは有意水準より小さい確率で起きるので帰無仮説を棄却する。コインは偏っていると考える。

単純化した問題設定

以下では $s = 0$ が帰無仮説、$s = 1$ が対立仮説に対応しているとする。
検定関数 $E\colon A \to \{0,1\}$ は $E(x) = 0$ のとき帰無仮説を採択、$E(x) = 1$ のとき帰無仮説を棄却することを表す。以下では確率的な検定関数 $E\colon A \to [0,1]$ を考える。$E(x) \in [0,1]$ は帰無仮説を棄却する確率を表す。(一般的には帰無仮説も対立仮説も分布の集合となるが、ここではそれぞれ一つの分布であることに注意する)

問題

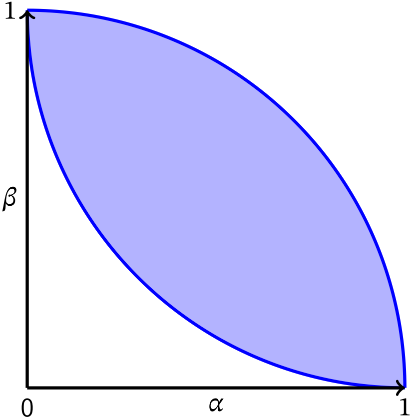
尤度関数 $\Pr(X = x \mid S = s)$ が既知である。与えられた $x \in A$ から $s \in \{0,1\}$ を推定したい。検定関数 $E\colon A \to [0,1]$ の第一種誤り確率 (有意水準) は $\alpha_E \coloneqq \mathbb{E}\left[E(X) \mid S = 0\right]$、第二種誤り確率は $\beta_E \coloneqq 1 - \mathbb{E}\left[E(X) \mid S = 1\right]$ で定義される。

第一種誤り確率を小さくしたければ $E(x) = 0$ とすればよいし第二種誤り確率を小さくしたければ $E(x) = 1$ とすればよい。

Definition (最強力検定)

検定関数 $E\colon A \to [0,1]$ が有意水準 $\alpha \in [0,1]$ の最強力検定 $\overset{\text{def}}{\iff} \alpha_E = \alpha$ であり、任意の $F\colon A \to [0,1]$ について $\alpha_F \leq \alpha$ ならば $\beta_F \geq \beta_E$ が成り立つ。

二種類の誤り確率



実現可能な (α, β) の集合 $\{(\alpha_E, \beta_E) \mid E : A \to [0,1]\}$ は凸集合である。
二つの検定関数 E_1 と E_2 の確率混合 $\lambda E_1(x) + (1-\lambda)E_2(x)$ はやはり検定関数で誤り確率も凸結合したものになる。

(α, β) が実現可能 $\iff$ $(1-\alpha, 1-\beta)$ も実現可能
$1 - E(x)$ を考えればよい。

尤度比検定

事前分布が分かっている場合、

$$E_{\mathrm{MAP}}(x) = \begin{cases} 0 & \text{if } \frac{\Pr(X=x|S=0)}{\Pr(X=x|S=1)} \geq \frac{\Pr(S=1)}{\Pr(S=0)} \\ 1 & \text{otherwise.} \end{cases}$$

が最適な検定関数となる。

一般的に $\eta > 0$, $\kappa \in [0, 1]$ について

$$E(x) = \begin{cases} 0 & \text{if } \frac{\Pr(X=x|S=0)}{\Pr(X=x|S=1)} > \eta \\ 1 & \text{if } \frac{\Pr(X=x|S=0)}{\Pr(X=x|S=1)} < \eta \\ \kappa & \text{otherwise.} \end{cases}$$

という形の検定関数を用いた仮説検定を尤度比検定という。

ネイマン・ピアソンの補題

Lemma (ネイマン・ピアソンの補題)

任意の尤度比検定は最強力検定である。*(逆も成り立つ)*。

Proof.

尤度比検定 E における尤度比の閾値を $\eta > 0$ とする。

$$p_0(x) \coloneqq \Pr(X = x \mid S = 0) \qquad p_1(x) \coloneqq \Pr(X = x \mid S = 1)$$

とする。任意の $F \colon A \to [0,1]$ について

$$\begin{aligned}
&(F(x) - E(x))(p_0(x) - \eta p_1(x)) \geq 0 \\
\iff\ & F(x)p_0(x) - E(x)p_0(x) - \eta F(x)p_1(x) + \eta E(x)p_1(x) \geq 0 \\
\implies\ & \alpha_F - \alpha_E - \eta(1 - \beta_F) + \eta(1 - \beta_E) \geq 0 \\
\iff\ & (\alpha_F - \alpha_E) + \eta(\beta_F - \beta_E) \geq 0.
\end{aligned}$$

よって $\alpha_F \leq \alpha_E$ ならば $\beta_F \geq \beta_E$ である。 $\square$

N回独立にサンプルする場合の尤度比検定

$$\Pr(\mathbf{X} = \mathbf{x} \mid S = s) = \prod_{k=1}^{N} \Pr(X_k = x_k \mid S = s)$$

$$\begin{aligned}
\log \frac{\Pr(\mathbf{X} = \mathbf{x} \mid S = 0)}{\Pr(\mathbf{X} = \mathbf{x} \mid S = 1)} &= \log \prod_{k=1}^{N} \frac{\Pr(X_k = x_k \mid S = 0)}{\Pr(X_k = x_k \mid S = 1)} \\
&= \sum_{k=1}^{N} \log \frac{\Pr(X_k = x_k \mid S = 0)}{\Pr(X_k = x_k \mid S = 1)} \\
&= N \sum_{x \in A} \hat{P}_{\mathbf{x}}(x) \log \frac{\Pr(X = x \mid S = 0)}{\Pr(X = x \mid S = 1)} \\
&= N \sum_{x \in A} \hat{P}_{\mathbf{x}}(x) \log \frac{p_0(x)\hat{P}_{\mathbf{x}}(x)}{p_1(x)\hat{P}_{\mathbf{x}}(x)} \\
&= N\left(D\left(\hat{P}_{\mathbf{x}} \,\|\, p_1\right) - D\left(\hat{P}_{\mathbf{x}} \,\|\, p_0\right)\right)
\end{aligned}$$

漸近的な誤り確率 1/3

Theorem (スタインの補題)

$D(p_0 \| p_1) < \infty$ とする。$\epsilon \in (0,1)$ について

$$\beta_N^{\epsilon} := \min_{\substack{E\colon A^N \to [0,1] \\ \alpha_E \le \epsilon}} \beta_E$$

とすると、任意の $\epsilon \in (0,1)$ について

$$\lim_{N\to\infty} \frac{1}{N} \log \beta_N^{\epsilon} = -D(p_0 \| p_1) .$$

上界の証明.

E を尤度比検定関数で $\kappa = 1$ とする。

$$\alpha_E = \Pr\left(\frac{p_0(\mathbf{X})}{p_1(\mathbf{X})} \le \eta \mid S = 0 \right)$$

$$\beta_E = \Pr\left(\frac{p_0(\mathbf{X})}{p_1(\mathbf{X})} > \eta \mid S = 1 \right)$$

□

漸近的な誤り確率 2/3

$$\alpha_E = \Pr\left(\frac{1}{N}\sum_{k=1}^{N}\log\frac{p_0(X_k)}{p_1(X_k)} \le \frac{1}{N}\log\eta \mid S=0\right)$$

$$\beta_E = \Pr\left(\frac{1}{N}\sum_{k=1}^{N}\log\frac{p_0(X_k)}{p_1(X_k)} > \frac{1}{N}\log\eta \mid S=1\right)$$

大数の弱法則やクラメールの定理により $\frac{1}{N}\sum_{k=1}^{N}\log\frac{p_0(X_k)}{p_1(X_k)}$ は

$$D(p_0 \,\|\, p_1) = \sum_x p_0(x)\log\frac{p_0(x)}{p_1(x)} \qquad \text{if } S=0$$

$$-D(p_1 \,\|\, p_0) = \sum_x p_1(x)\log\frac{p_0(x)}{p_1(x)} \qquad \text{if } S=1$$

に集中する。よって E_N において $\frac{1}{N}\log\eta = D(p_0 \,\|\, p_1) - \delta$ とすると $\alpha_{E_N} \to 0$ である。

漸近的な誤り確率 3/3

$$
\begin{aligned}
\beta_{E_N} &= \Pr\left(\frac{1}{N}\sum_{k=1}^{N}\log\frac{p_0(X_k)}{p_1(X_k)} > D\left(p_0 \,\|\, p_1\right) - \delta \mid S = 1\right) \\
&= \Pr\left(D\left(\hat{P}_{\mathbf{X}} \,\|\, p_1\right) - D\left(\hat{P}_{\mathbf{X}} \,\|\, p_0\right) > D\left(p_0 \,\|\, p_1\right) - \delta \mid S = 1\right) \\
&\le \Pr\left(D\left(\hat{P}_{\mathbf{X}} \,\|\, p_1\right) > D\left(p_0 \,\|\, p_1\right) - \delta \mid S = 1\right)
\end{aligned}
$$

よってサノフの定理より

$$
\limsup_{N\to\infty}\frac{1}{N}\log\beta_{E_N} \le -D\left(p_0 \,\|\, p_1\right) + \delta
$$

任意の $\delta > 0$ について成り立つので $\delta \to 0$ 極限で
$\limsup_{N\to\infty}\frac{1}{N}\log\beta_{E_N} \le -D\left(p_0 \,\|\, p_1\right)$.

連続確率変数の場合

- $x \in A$: データ、知っているもの、$A \subseteq \mathbb{R}$ とか $A \subseteq \mathbb{R}^n$ とか。

- $s \in B$: パラメータ、知りたいもの、B は有限集合とする

- $\Pr(S = s)$: 事前分布

- $p(x \mid s)$: 尤度 (条件付き確率密度関数)

- $p(s \mid x) \coloneqq p(x \mid s) \Pr(S = s)/p(x)$: 事後確率

$$\Pr(X \in B,\ S = s) = \int_B p(x \mid s) \Pr(S = s) \mathrm{d}x$$

$$p(x) = \sum_{s \in B} p(x \mid s) \Pr(S = s)$$

連続確率変数の場合のMAP,ML推定

最大事後確率推定 (Maximum a posteriori probability estimation)

$$E_{\mathrm{MAP}}(x) \coloneqq \operatorname*{arg\,max}_{s \in B} p(s \mid x)$$

平均誤り確率を最小化する。事前分布を知らないと使えない。

最尤推定 (Maximum likelihood estimation)

$$E_{\mathrm{ML}}(x) \coloneqq \operatorname*{arg\,max}_{s \in B} p(x \mid s)$$

平均誤り確率については何も言えない。事前分布を知らなくても使える。

事前分布が一様分布のとき、最尤推定と最大事後確率推定は等しい。

連続確率変数の場合の尤度比検定

事前分布が分かっている場合、

$$E_{\mathrm{MAP}}(x) = \begin{cases} 0 & \text{if } \frac{p(x|\theta_0)}{p(x|\theta_1)} \geq \frac{\Pr(S=\theta_1)}{\Pr(S=\theta_0)} \\ 1 & \text{otherwise.} \end{cases}$$

が最適な検定関数となる。

一般的に $\eta > 0,\ \kappa \in [0,1]$ について

$$E(x) = \begin{cases} 0 & \text{if } \frac{p(x|\theta_0)}{p(x|\theta_1)} > \eta \\ 1 & \text{if } \frac{p(x|\theta_0)}{p(x|\theta_1)} < \eta \\ \kappa & \text{otherwise.} \end{cases}$$

という形の検定関数を用いた仮説検定を尤度比検定という。

ネイマン・ピアソンの補題も成り立つ。

分散が同じで既知な正規分布の尤度比検定

$$p(x \mid \theta_0) = \frac{1}{\sqrt{2\pi\sigma^2}} \mathrm{e}^{-\frac{(x-\theta_0)^2}{2\sigma^2}}$$

$$p(x \mid \theta_1) = \frac{1}{\sqrt{2\pi\sigma^2}} \mathrm{e}^{-\frac{(x-\theta_1)^2}{2\sigma^2}}$$

のとき、

$$\frac{p(x \mid \theta_0)}{p(x \mid \theta_1)} = \mathrm{e}^{\frac{2(\theta_0-\theta_1)x+\theta_1^2-\theta_0^2}{2\sigma^2}}$$

$$\log \frac{p(\mathbf{x} \mid \theta_0)}{p(\mathbf{x} \mid \theta_1)} = N\frac{\theta_1^2 - \theta_0^2}{2\sigma^2} + \frac{\theta_0 - \theta_1}{\sigma^2} \sum_{k=1}^{N} x_k$$

なので $\sum_k x_k$ の大きさを見て検定することになる。

課題

$A = \{0, 1\}$, $B = \left\{p_0 = \frac{1}{2}, p_1 = \frac{1}{3}\right\}$ とし、$k \in \{0, 1\}$ について

$$\Pr(X = 0 \mid S = p_k) = 1 - p_k \qquad \Pr(X = 1 \mid S = p_k) = p_k$$

とする。このとき、$\eta > 0, \kappa \in [0, 1]$ について、以下の尤度比検定関数 $E(x)$ で仮説検定することを考える。

$$E(x) = \begin{cases} 0 & \text{if } \frac{\Pr(X=x|S=p_0)}{\Pr(X=x|S=p_1)} > \eta \\ 1 & \text{if } \frac{\Pr(X=x|S=p_0)}{\Pr(X=x|S=p_1)} < \eta \\ \kappa & \text{otherwise.} \end{cases}$$

以下の問に答えよ。

1. $\eta = 1$ のとき α_E, β_E をもとめよ。
2. 一般の $\eta > 0$ と $\kappa \in [0, 1]$ について、誤り確率 α_E, β_E をもとめよ。η の値で場合分けしてもとめること。
3. 任意の検定関数 $E \colon A \to [0, 1]$ で実現可能な (α_E, β_E) の範囲を図示せよ。